Using 19F NMR and two-level factorial design to explore thiol-fluoride substitution in hexafluorobenzene and its application in peptide stapling and cyclisation

Received: 30 April 2020 | Revised: 21 June 2020 | Accepted: 22 June 2020

DOI: 10.1002/pep2.24182

FULL PAPER

PeptideScience WILEY

Using ^{19}F NMR and two-level factorial design to explore thiol-fluoride substitution in hexafluorobenzene and its application in peptide stapling and cyclisation

Paolo Dognini[1] | Patrick M. Killoran[2] | George S. Hanson[1,3] | Lewis Halsall[1] | Talhat Chaudhry[1] | Zasharatul Islam[1] | Francesca Giuntini[1] | Christopher R. Coxon[3]

[1]School of Pharmacy and Biomolecular Sciences, Byrom Street Campus, Liverpool John Moores University, Liverpool, UK

[2]Division of Structural Biology (STRUBI), University of Oxford, Rutherford Appleton Laboratory, Harwell Science and Innovation Campus, Harwell, UK

[3]Institute of Chemical Sciences, School of Engineering and Physical Sciences, Heriot-Watt University, Edinburgh, UK

Correspondence
Christopher R. Coxon, Institute of Chemical Sciences, School of Engineering and Physical Sciences, Heriot-Watt University, Edinburgh EH14 4AS, UK.
Email: c.coxon@hw.ac.uk

Funding information
Engineering and Physical Sciences Research Council, Grant/Award Number: EP/R020299/1; H2020 Marie Skłodowska-Curie Actions, Grant/Award Number: 801604

Abstract

Hexafluorobenzene undergoes 1,4-selective thiol-fluoride disubstitution and is an attractive disulfide crosslinking reagent for peptide cyclisation and stapling. Little attention has been directed toward understanding the scope of this reaction. Traditional reaction optimisation relies on a one-variable-at-a-time approach, which can exclude important combined effects of reaction variables. This study initially explored base and solvent effects to inform a subsequent two-level factorial design approach to understand how to control the reactivity and product selectivity in a model reaction of hexafluorobenzene. We describe new conditions that selectively afford higher order substitution products for example, 1,2,4,5-tetrasubstitution, making hexafluorobenzene a possible suitable scaffold for future branched or multicyclic peptide systems. Moreover, our new conditions provide an improved rapid (<1 minute) and selective peptide disulfide stapling and cyclisation approach under peptide-compatible conditions.

KEYWORDS

^{19}F NMR, design of experiments, factorial design, hexafluorobenzene, peptide stapling

1 | INTRODUCTION

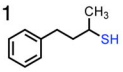
Hexafluorobenzene (HFB) is the most elementary perfluoroaromatic system, prepared by reaction of hexachlorobenzene at high temperature and pressure with potassium fluoride.[1] It has been applied to investigate tissue oxygenation *in vivo*,[2] but it can also be used as a solvent in organic reactions,[3] in ^{1}H/^{13}C NMR, UV and IR spectroscopy and as a reference standard for ^{19}F NMR. The presence of six inductively electron withdrawing fluorine atoms (the F atom has the highest Pauling scale electronegativity value in the periodic table), makes HFB highly electrophilic and promotes the ring reactivity to be dominated by nucleophilic aromatic substitution (S_NAr). S_NAr on polyfluorinated arenes is a valuable method for constructing highly functionalised heterocyclic aromatic molecules. Fluoride displacement takes place with a range of nucleophiles,[4] including thiols (R-SH). The S_NAr reaction on electron-withdrawing group activated rings, usually performed in a basic medium, generally proceeds via an addition-elimination mechanism with the intermediate formation of a reactive negatively charged adduct between the arene and the nucleophile (Meisenheimer complex).[5] When HFB reacts with 2 M equivalents of a thiol nucleophile, the anion stabilization of the first thiol increases the rate of the second substitution affording exclusively disubstitution products with a 1,4-regiosubstitution pattern. Subsequent substitutions, generally observed with higher amounts of thiols and

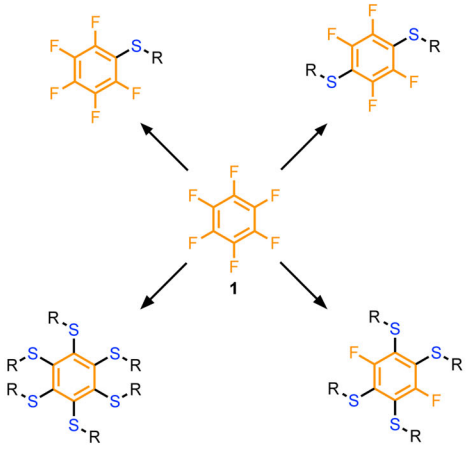
increasingly vigorous reaction conditions, follow the same regioselectivity, resulting in 1,2,4,5-tetrasubstitution and possible 1,2,3,4,5,6-hexasubstitution. Mono-, tri- and penta-substitution products are rarely isolated in significant quantities. In principle, the extent of substitution can be ordered by careful control of the reagents nucleophilicity and the reaction conditions and is summarized in Figure 1.[6–20] In general, there are few approaches reported to selectively afford mono-substitution of HFB, while disubstitution is generally obtained under mild conditions at room temperature. Stronger bases, longer reaction times, higher temperatures and a greater concentration of thiols can orient the reaction toward tetra-substitution while hexa-substitution can be achieved with even harsher conditions and a larger excess of thiols. Aromatic thiolates (i.e., thiophenol or derivatives) react more readily than aliphatic thiolates and this difference influences the reaction conditions.

HFB has broad applications as a molecular scaffold for cysteine-containing peptide stapling and multicyclisation. In the presence of a deprotected cysteine-containing peptide, HFB undergoes a 1,4-disubstitution with no observable higher order substitution products (Scheme 1). Such transformations are typically performed employing DMF as solvent and either TRIS-base, Et_3N or DIPEA as base (Table 1).[15,21–28] In the work from Zhang *et al*,[22] the reaction is promoted by a glutathione *S*-transferase (GST) enzyme that, after TCEP-mediated deprotection of the Cys (StBu), catalyzes the macrocyclisation. However, this is only applicable when the peptide sequence contains a γ-Glu-Cys-Gly (GSH) motif. Examples of the same reaction with different base-solvent combinations or alternative methodologies are not reported.

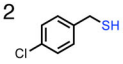
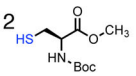
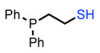
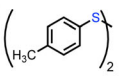
The remaining four fluorine atoms are conceptually also available for substitution, making this an excellent scaffold for stepwise

1-monosubstitution

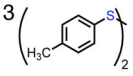
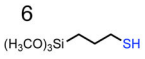
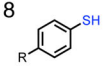
Conditions	*Thiol (eq.)*	*Ref*
K_2CO_3, DMF, RT, 4.5 h	1 	6
THF, -40 °C	1 	7



1,4-disubstitution

Conditions	*Thiol (eq.)*	*Ref*
NaH, DMF, RT, overnight	2 	13
TRIS base, DMF, RT, 3-5 h	2 	14
TRIS base, DMF, RT, 4.5 h	2 	15
Cs_2CO_3, NBu_4Cl, THF, RT, 1-3 days	2 	16
Cs_2CO_3, NBu_4Cl, THF, RT, 18 h	2 	17
Et_2O, RT, 2 days	2 	18
$(RhH(PPH_3)_4$, dppBz, PPh_3, PhCl, 80 °C, 4 h	1 	19

1,2,3,4,5,6-hexasubstitution

Conditions	*Thiol (eq.)*	*Ref*
$(RhH(PPH_3)_4$, dppBz, PPh_3 , PhCl, 80 °C, 48 h	3 	19
HMPA, 0 °C, 1.5 h	10 	9
NaH, THF, RT, 24 h	6 	10
Cs_2CO_3, DMI, 35-50 °C, 3-6 days	10 15 	11
NaH, DMI, RT, 6 h - 3 days	18 	12

1,2,4,5-tetrasubstitution

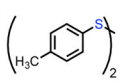
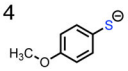
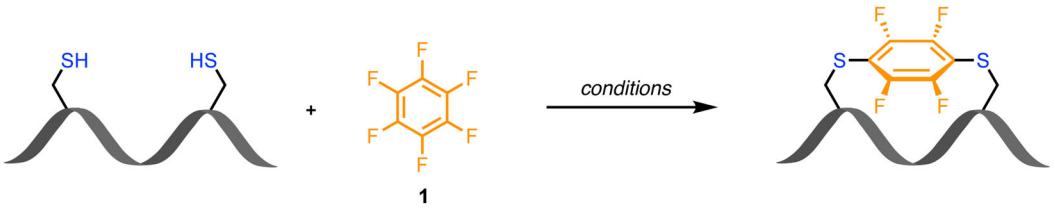
Conditions	*Thiol (eq.)*	*Ref*
DMI/EtOH (1:1), 80 °C, 75 h	4 	20
Cs_2CO_3, NBu_4Cl, THF, RT, 3-14 days	4 	16
$(RhH(PPH_3)_4$, dppBz, PPh_3, PhCl, 80 °C, 12 h	2 	19
DMI, 40 °C, 72 h	4 	8
HMPA, 0 °C, 0.1 h	4 	9

FIGURE 1 General approaches to selectively access mono-, di-, tetra- and hexa- substituted fluorobenzenes



SCHEME 1 General reaction scheme for a S_NAr on HFB with two-cysteine-containing peptides

TABLE 1 General stapling and cyclisation approaches that exploits the reaction between di-cysteine peptides and HFB

Conditions	Peptides	Ref
TRIS-base, DMF, rt, 4.5 hours	YCGGGCAL, YCERSCNMK, ITFCDLLCYYGKKK, CNLLCEAKKLNDAQAPK	[15]
TRIS-base, DMF, rt, 2 hours	From *i*,*i* + 1 [IKFTNGLCCLYESKR] to *i*,*i* + 14 [CIKFTNGLLYESKRC]	[21]
1. TRIS-base, DMF, rt, 30 minutes 2. GST, TCEP, 0.1 M phosphate buffer (pH 8), rt, 5 minutes	γ-EC(StBu)G(GLKAG)$_x$C X = 2; 3; 4	[22]
TRIS-base, DMF, rt, 18 hours	CDEETGEC, CEETGC, CDPETGEC, CLDPETGEFC	[23]
TRIS-base, DMF, rt	$HX_1EGCX_2TSCX_3X_4$, $HX_1EGTX_2TCDX_3X_4C$	[24]
Et_3N, DMSO/DMF (1:1), rt, 14 hours	CGNKRTRGC	[25]
DIPEA, DMF, rt, 4 hours	YCGGGCAL, FKACGKGCA	[26]
DIPEA, DMF, rt, 4 hours	LTFCHYWCQLTS	[27]
TRIS-base, TCEP, DMF, rt, 2 hours	WGKGCGKGUGKGCW	[28]

multicyclisation or poststapling introduction of additional functionality. The literature indicates that further substitution is possible using nonpeptide thiols and typically requires either higher temperature, longer reaction times and/or stronger inorganic bases.

Factorial design (FD) is a powerful experimental design approach that, with a small number of experiments, provides high-quality information in the whole experimental domain. The alternative One-Variable-At-a-Time (OVAT) approach gives very specific information only for the conditions in which the experiments have been performed. However, a FD reflects the interactions between controllable variables (factors), while the OVAT would only be valid if the variables are totally independent from each other.[29] Therefore, FD is advantageous with respect to time, money and relevancy of the results and can be usefully applied in the chemistry field. Time, temperature, volume of solvent(s) and amount of reagent(s) are just a few variables that play a role within, or may affect the outcome of, a standard laboratory reaction. Consequently, they are factors frequently considered during optimisation processes and their mutual interaction is highly likely. FD allows these to be studied individually, in addition to the interactions between them to be observed.

The fluorine atoms of fluorobenzenes provide a diagnostic reporter for product formation by ^{19}F NMR. ^{19}F NMR is a fast, easily interpretable, quantifiable and reproducible reaction monitoring and optimisation tool. The number of substituted fluorine atoms on the aromatic ring is a direct indicator of the substitution degree on the molecule. Fluorine signal integration is quantitative and, therefore, allows for accurate quantification of the relative amount of each product. Having no background fluorine, ^{19}F NMR allows for easier characterization of chemicals compared to ^{1}H NMR. In this case, only LC-MS analysis can be as much informative as NMR analysis is, but solvent usage and long run times add value to NMR practicality.

Exploration of the S_NAr reaction between HFB and *N*-acetyl cysteine will help peptide chemists to understand what drives the level of HFB substitution when reacted with a thiol. Here, after initial screenings of different base-solvent combinations, we apply a FD approach coupled with ^{19}F NMR to investigate the parameters that predominantly affect the product outcomes in the reaction of HFB with *N*-acetyl cysteine as a model for upcoming peptide studies. This work also introduces three new sets of conditions to execute the same reaction for peptide stapling and cyclisation.

2 | METHODS AND MATERIALS

2.1 | General points

All the reagents were purchased from commercial suppliers and used directly as indicated in the appropriate experimental procedures. All Fmoc L-amino acids with standard side-chain protecting groups (Fmoc-Ala-OH, Fmoc-Cys(Trt)-OH, Fmoc-Gly-OH, Fmoc-His(Trt)-OH, Fmoc-Leu-OH, Fmoc-Gln(Trt)-OH, Fmoc-Ser(tBu)-OH, Fmoc-Thr(tBu)-OH, Fmoc-Tyr(tBu)-OH), Rink Amide Pro-Tide resin and Oxyma Pure were purchased from CEM UK Ltd (Buckingham, UK). TFA and DIC were purchased from Fluorochem UK Ltd (Hadfield, UK). DMF was purchased from Fisher Scientific UK Ltd (Loughborough, UK). DIPEA, piperidine, HFB, triisopropylsilane, DMSO, MeCN, cesium carbonate, potassium carbonate, THF, DBU, propylene carbonate (PC), Et_2O were purchased from Sigma Aldrich Company Ltd (Gillingham, UK).

2.2 | ^{19}F NMR analysis

All proton-decoupled ^{19}F NMR spectra where recorded on a Bruker Avance 600 MHz NMR, operating at 564.7 MHz at 300 K

in high-resolution NMR tubes with 0.7 mL of each sample. Signals were observed as singlets for the products of interest, which were assigned based on known ^{19}F NMR resonance data.[15,16] Data analysis (integration, peak-picking) was carried-out using TopSpin 3.1, Bruker UK Ltd (Coventry, UK). All spectra are provided in Figures S5-S35).

2.3 | Liquid chromatography-mass spectrometry (high resolution) analysis

Characterization of crude peptides and reaction products were conducted using an Agilent 1260 Infinity II LC system with Agilent 6530 Accurate-Mass QToF spectrometer, equipped with an Agilent ZORBAX SB-C18 Stable Bond Analytical (5 μm particle size, 4.6 × 150 mm) from Agilent Technologies UK Limited (Cheadle, UK) with a binary eluent system comprising MeCN/H_2O (25 minutes gradient: 1%-99% with 0.1% formic acid) as mobile phases. MS grade solvents (MeCN, formic acid) were purchased from Fisher Scientific and the employed ultrapure H_2O was purified with a Milli-Q Reference Water Purification System. Electrospray ionization mass spectrometry was conducted in positive ion mode (m/z range: 50-3200) using a fragmentor voltage of 150 V, gas temperature of 325 °C (flow 10 L/min) and sheath gas temperature of 400 °C (flow 11 L/min).

2.4 | Model reactions - OVAT approach

Reactions were performed in 14 mL screw top vials and stirred using stirrer bars and magnetic stirrer-plates. *N*-acetyl L-cysteine (113 mg, 0.692 mmol, 4 eq.), the required solvent (5 mL), the required base (3.46 mmol, 20 eq.) and HFB (20 μL, 0.173 mmol, 1 eq.) were added to the reaction vessel and mixtures were stirred for 4 hours at 21 °C. The reaction outcomes were analyzed with ^{19}F NMR in order to measure the percentage of unreacted HFB and various substitution products. The above general procedure was applied to combinations of solvents (THF, MeCN, DMF, DMSO and PC) and bases (DIPEA, cesium carbonate and DBU) as indicated. *1,4-disubstitution product* (**3**): HRMS $[M + H]^+$ *m/z* for $[C_{16}H_{17}F_4N_2O_6S_2]^+$ calculated 473.03859, found 473.04691; *1,2,4,5-tetrasubstitution product* (**4**): HRMS $[M + H]^+$ m/z for $[C_{26}H_{33}F_2N_4O_{12}S_4]^+$ calculated 759.08676, found 759.09531. See Supporting Information for complete LCMS characterization (Figure S3).

2.5 | Factorial design

A two-level FD space was generated in Minitab (version 19) to assess the influence of 5 factors: thiol (mol. eq.), base (mol. eq.), reaction time (h), temperature (°C) and reaction volume (mL) (Table 2). High (maximum) and low (minimum) values for each factor chosen based on earlier observations, were combined in an array of 32 individual experiments (See Supporting Information for complete details and raw data). Each reaction was performed in a 10 mL fritted syringe containing a magnetic stirrer bar. *N*-acetyl cysteine (113 mg, 4 eq. or 282 mg, 10 eq.), Cs_2CO_3 (450 mg, 8 eq. or 1129 mg 20 eq.) were added to the reaction vessel, followed by a solution of HFB (20 μL, 1 eq.) in DMSO (5 mL or 10 mL), leaving no headspace. The reaction vessel was then capped and stirred for 4 or 168 hours at the appropriate temperature (21 °C or 50 °C). For experiments that were performed at an elevated temperature of 50 °C the vessel was stirred in a water bath at 50 °C, maintained by using a temperature probe. At the completion of the reaction, the contents of the syringe were filtered through the frit and 700 μL was used for ^{19}F NMR analysis (data provided in Figures S5-S35). Statistical analysis, including ANOVA, factorial regression and main effects plots were generated in Minitab.

TABLE 2 Factorial design space - high and low settings for each factor

Controllable factor	Low setting	High setting
Time	4 hours	168 hours
Solvent volume	5 mL	10 mL
Temperature	21 °C	50 °C
Mol. equation *N*-acetyl cysteine	4	10
Mol. eq. Cs_2CO_3	8	20
Solvent	DMSO	

2.6 | Solid phase peptide synthesis

Each linear di-cysteine peptide sequence was prepared using automated Fmoc-SPPS methods on a CEM Liberty Blue microwave-assisted peptide synthesizer. Solid-phase synthesis was conducted using Rink amide Pro-Tide resin (180 mg, 0.56 mmol/g loading; 0.1 mmol), employing the required Fmoc amino acids (0.2 M in DMF, 5 eq.); DIC (1 M stock solution in DMF; 10 eq.) as activator, Oxyma Pure (1 M stock solution, 5 eq.) as racemisation suppressor, and piperidine (20% *v/v* in DMF; 587 eq., 4 mL) as deprotection agent. Standard coupling procedures employed single coupling of each amino acid (2.5 minutes, 90 °C) and Fmoc-deprotection (2 minutes, 90 °C). Racemisation-prone amino acids bearing thermally-sensitive protecting groups, for example, Fmoc-Cys(Trt)-OH were coupled under milder conditions (10 minutes, 50 °C). Following on-resin synthesis of the appropriate sequence, the resin was transferred in 10 mL syringes with frits and shrank with Et_2O. Finally, peptides were cleaved from the resin as the C-terminal amide by treatment with a cleavage cocktail comprising TFA, TIPS and H_2O (8:1:1 *v/v*) with shaking at 21 °C for 4 hours. Peptides were precipitated from cleavage solutions by dropwise addition into cold Et_2O followed by centrifugation. The resulting pellet was successively suspended in cold Et_2O and centrifuged twice further. The solids obtained after Et_2O removal and its complete evaporation were analyzed by LCMS. For LCMS analysis, the peptide was dissolved in H_2O, MeCN or MeOH with 0.1% formic

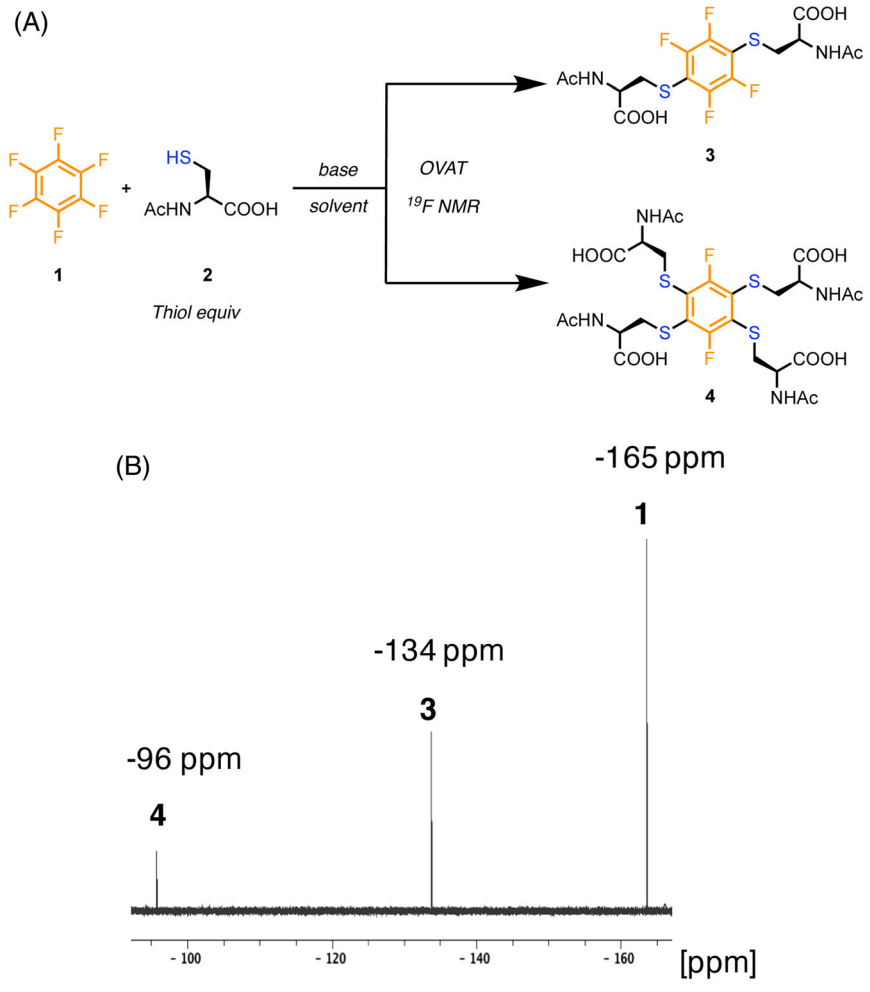
acid *v/v* (or a mixture of these solvents, depending on the solubility) and analyzed by LCMS (as above). The crude peptides were used for successive reactions without further chromatographic purification. Peptide (**5**) H-YCGGGCAL-NH_2: HRMS $[M + H]^+$ *m/z* for $[C_{30}H_{48}N_9O_9S_2]^+$ calculated 742.29382, found 742.30281; *SPACE peptide* (**7**) Acetyl-ACTGSTQHQCG-NH_2: MS $[M + H]^+$ *m/z* for $[C_{42}H_{67}N_{16}O_{17}S_2]^+$ calculated 1133.43898, found 1133.09. See Supporting Information for complete LCMS characterization—Figures S39, S43.

2.7 | General procedure for peptide stapling in solution

The crude peptide (1 eq.) and any other solid reagent (e.g., Cs_2CO_3, 20 eq. or TCEP 2 eq.) were weighed and solubilized in small vials (from 1.5 to 3 mL, 1.5 mL of solvent for 50 mg of peptide) with magnetic stirring. Any other liquid reagent (e.g., DBU, 20 eq.) and, last, HFB (10 eq. or 0.5 eq. as required) were added to a sealed vial that was stirred at 21 °C for 24 hours, with sampling at intermediate timepoints. Twenty micro liter of reaction mixture were sampled and analyzed by NMR and LCMS after dilution with 1 mL H_2O + 0.1% TFA. Peptide (**6**) from DMSO/Cs_2CO_3 reaction: HRMS $[M + H]^+$ *m/z* for $[C_{36}H_{46}F_4N_9O_9S_2]^+$ calculated 888.27178, found 888.27962; peptide (**6**) from MeCN/DBU reaction: HRMS $[M + H]^+$ *m/z* for $[C_{36}H_{46}F_4N_9O_9S_2]^+$ calculated 888.27178, found 888.27859. See Supporting Information for complete LCMS characterization—Figures S40-S42.

2.8 | General procedure for peptide stapling on resin

After SPPS, the mass of dry resin corresponding to 1 eq. of peptide was weighed directly in an empty 10 mL fritted-syringe and swollen for 5 to 10 minutes in DMF. Trityl protected cysteine residues were selectively deprotected on resin with a cleavage solution comprising DCM:TFA:TIPS (35:4:1, 2 mL/10 mg of resin), for 1 hour at 21 °C.[30] The cleavage solution was discarded, and the resin washed successively with DCM and DMF. A solution of HFB (10 eq.) and DIPEA (0.4 M in DMF; 1 mL/10 mg of resin) was drawn up into the syringe containing the swollen peptide resin and shaken at 21 °C for 18 hours. Subsequently, the reaction mixture was gently evaporated under a stream of nitrogen gas, the resin washed with DCM and DMF and shrunk with Et_2O. The stapled/cyclic product was finally cleaved



SCHEME 2 A, Model reaction between HFB and *N*-acetyl cysteine. B, Typical ^{19}F NMR for this model reaction, showing characteristic quantifiable resonances for substitution products. These chemical shifts are in accordance with previously reported literature assignments.[15,16]

TABLE 3 Summary of solvent properties and the effect of solvent-base combinations on product outcome

Solvent	Dielectric constant (ε)	Polarisability (α)	Viscosity at 25 °C (cP)	Boiling point (°C)
THF	7.6	7.94	0.48	66
MeCN	37	4.44	0.33	82
DMF	38	7.93	0.8	153
DMSO	47	8.03	1.99	189
PC	64	8.55	22.4	240

DIPEA	pK_a of conjugate acid (in H_2O) = 10.75 [35]			
Solvent	Unreacted	Monosubstituted	Disubstituted	Tetrasubstituted
THF	100	0	0	0
MeCN	92	7.34	0.66	0
DMF	41	22	37	0
DMSO	1	0	99	0
PC	20	5	75	0

DBU	pK_a of conjugate acid (in H_2O) = 11.5,[36] 11.9 [37]			
Solvent	Unreacted	Monosubstituted	Disubstituted	Tetrasubstituted
THF	68	0	32	0
MeCN	0	10	80	10
DMF	0	12	85	3
DMSO	0	0	60	40
PC	0	0	97	3

Cs_2CO_3	pK_a of conjugate acid (in H_2O) = 10.33 [38]			
Solvent	Unreacted	Monosubstituted	Disubstituted	Tetrasubstituted
THF[a]	n/a	0	100	0
DMF[b]	0	0	93	7
DMSO	0	0	63	37

Note: Darker colors provide a visual aid to indicate higher conversion.
[a]In precipitated solid only.
[b]In precipitated solid but no starting material remained in solution.

and analyzed following the procedure described above. Peptide (**6**): HRMS $[M + H]^+$ *m/z* for $[C_{36}H_{46}F_4N_9O_9S_2]^+$ calculated 888.27178, found 888.27915; *SPACE-TFB peptide* (**8**): MS $[M + H]^+$ *m/z* for $[C_{48}H_{67}F_4N_{16}O_{17}S_2]^+$ calculated 1279.41694, found 1279.14. See Supporting Information for complete LCMS characterization—Figure S44.

3 | RESULTS AND DISCUSSION

3.1 | Initial condition scoping—OVAT

The initial aim of this study was to explore how, by applying specific combinations of reaction conditions, it is possible to selectively obtain higher conversion to either di- or tetra-thiol-substituted fluorobenzenes, under mild and peptide-friendly conditions. A wide range of procedures have been reported to afford similar products (Figure 1), however, many of these are not well suited to peptides and amino acids. For example, the use of temperatures higher than 50 °C, metal catalysis and a large excess of thiols should ideally be avoided. In order to narrow the scope of conditions that could be applied to a FD space, the model reaction between N-acetyl cysteine and HFB was studied using an initial OVAT approach (Scheme 2, A). As HFB and the two major products both contain all equivalent fluorine nuclei, each exhibited a characteristic singlet at approximately −165, −134 and −96 ppm for unreacted HFB, disubstitution and tetrasubstitution products, respectively (Scheme 2B).[15,16] This allowed relative quantification of the main reaction products by signal integration and dividing by the number of fluorine nuclei, Figure S2.

Preliminary studies (see Figure S4) showed that disubstitution was obtained as the predominant product using the organic base DIPEA, even with increasing equivalents of thiol and time. However, the inorganic base K_2CO_3 generally encouraged a larger proportion of higher-order substitution products, particularly when using a slightly wasteful 10-fold excess of thiol. At least 2 mol. Equivalents of base (with respect to HFB) were required as 1 equivalent afforded a mixture of mainly unreacted, mono-substituted and a small proportion of disubstituted fluorobenzene, while in the absence of base no reaction

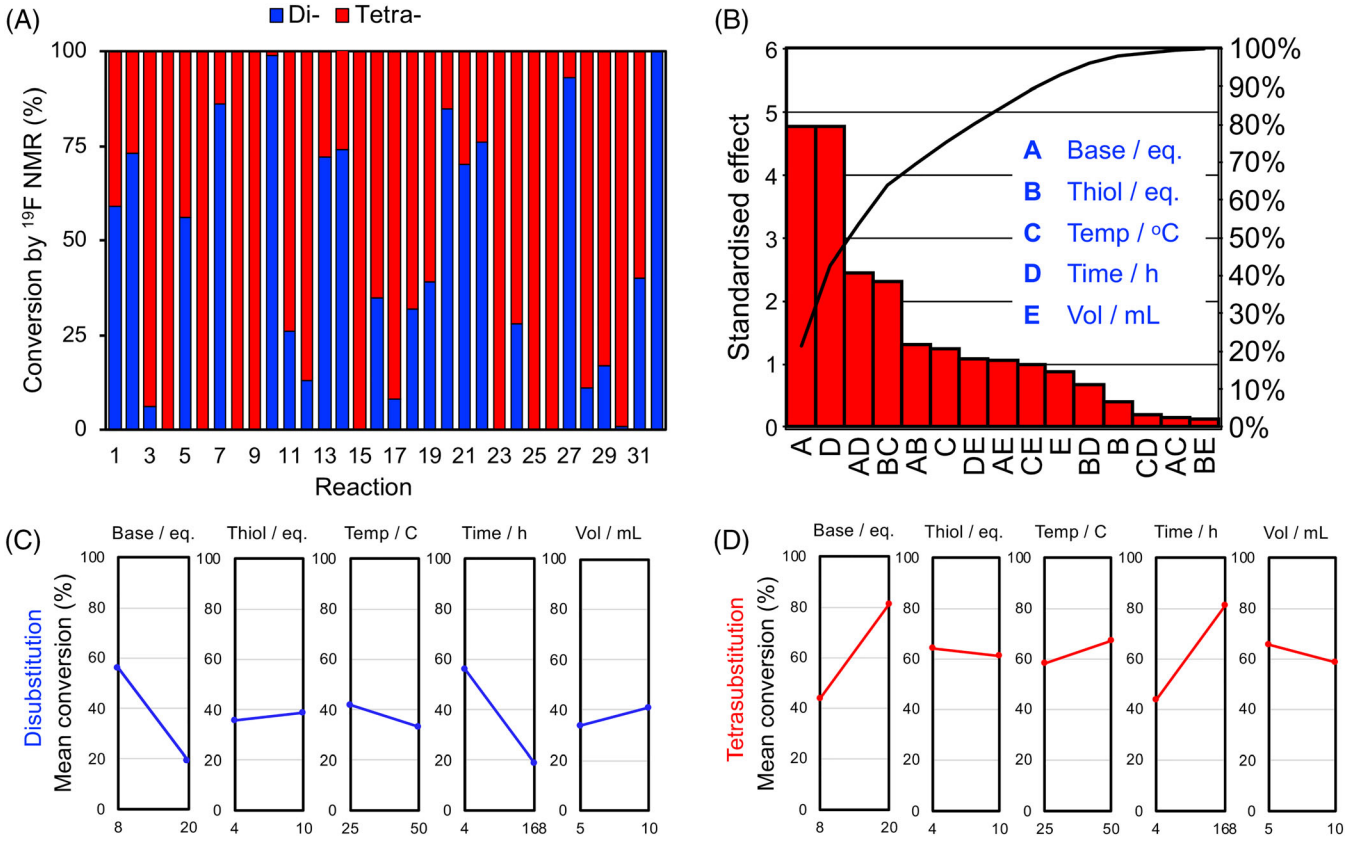


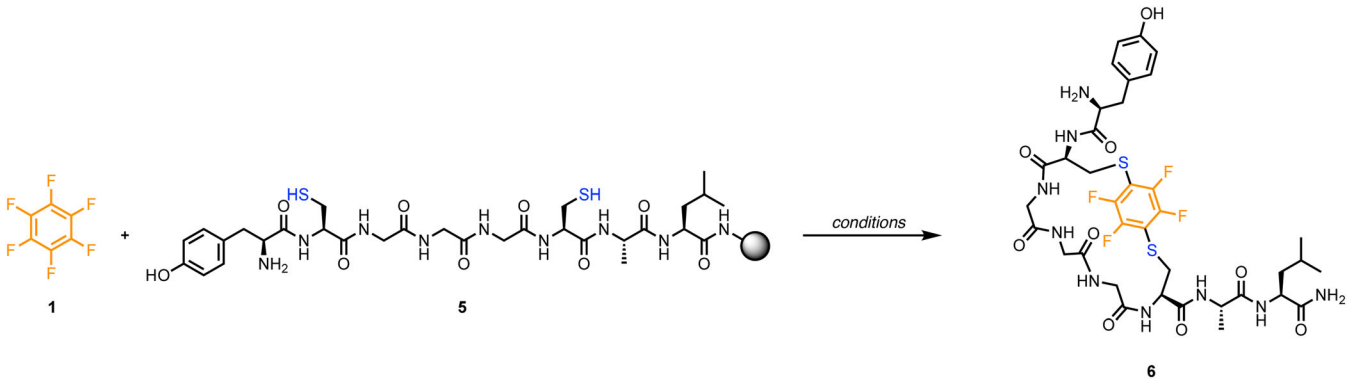
FIGURE 2 Two-Level factorial design analysis: A, Product distribution and relative % conversion by ^{19}F NMR. B, Pareto plot to indicate factors that influence reaction outcome. C, Main effects plot for 1,4-disubstitution (%), and D, 1,2,4,5-tetrasubstitution (%)

TABLE 4 Perfluoroaryl stapling of model peptides under new conditions

Conditions	C-terminus	Peptide sequences
Cs_2CO_3, DMSO, 21 °C, 5 minutes	H (free peptide)	Y**C**GGG**C**AL (**6**)
DBU, MeCN, 21 °C, 1 hour	H (free peptide)	Y**C**GGG**C**AL (**6**)
DIPEA, DMF, 21 °C, 18 hours	Resin	Y**C**GGG**C**AL (**6**), A**C**TGSTQHQ**C**G (SPACE peptide)

took place. Gentle heating (40 °C) afforded lower conversion when compared to the same reaction at RT, likely due to the relative volatility of HFB.

We have previously reported that careful selection of solvents can increase chemoselectivity in thiol-fluoride substitutions in reactions of peptide amines/thiols with pentafluoropyridine.[26,31] Subsequently, additional bases were studied in a range of polar aprotic solvents to try to gauge the importance of properties such as dielectric constant and polarizability (Table 3) upon promoting substitution at the fluorinated aromatic ring. Several solvents were compared, including DMSO, DMF, MeCN, THF and propylene carbonate, which has recently shown promise as a "green" alternative to DMF for peptide synthesis.[32] For each solvent, the bases DIPEA, DBU and cesium carbonate were probed. The reactions were performed in the required solvent (5 mL) using the required base (20 mol. eq.) and HFB (34.6 mM, 1 mol eq.), and the mixtures were stirred for 4 hours at 21 °C prior to ^{19}F NMR analysis. Using DIPEA as a base afforded little or no conversion to substituted products in acetonitrile and THF, while DMF provided a near-equal distribution of unreacted (41%), monosubstituted (22%) and disubstituted (37%) products. Disubstitution proceeded efficiently in both DMSO and propylene carbonate, despite the latter also presenting some unreacted starting material. The stronger organic base DBU generally seemed to promote higher conversions to substituted products (significant variation when switching from DIPEA to DBU in MeCN), following a similar solvent-trend as for DIPEA. However, of particular note in this case was the moderate production (40%) of the tetrasubstituted product in DMSO. Replacement of the organic bases with the inorganic base Cs_2CO_3 presented an additional challenge: the base was mostly insoluble in each solvent and most reactions afforded an additional precipitate. Redissolution and acidification of the precipitate confirmed by ^{19}F NMR that the substitution products were precipitated as presumed insoluble cesium salts. In THF the precipitate was entirely of the disubstituted product, however, it should be noted that this cannot be accurately quantified due to the presence of some unreacted starting material remaining in solution. In DMF, no unreacted material remained in solution and afforded a



SCHEME 3 Perfluoroaryl stapling of model peptides

precipitate as a mixture of disubstituted (major) and tetrasubstituted (minor) products. When DMSO was used as the solvent, all products remained in solution and afforded a similar product distribution as was the case for DBU in DMSO, with around 40% tetrasubstituted product. In summary, there appears to be an overall trend for increasing solvent dielectric and polarisability to afford higher conversion to disubstitution and in some cases (DMSO as solvent) the tetrasubstituted product. Therefore, the higher polarity solvent may be a better insulator of the charges of the thiolate nucleophile, and the Meisenheimer complex formed in the reaction.[33] The outlier to this trend is propylene carbonate which has the highest polarisability and dielectric properties of the solvents investigated, yet seems to be less efficient than DMSO in promoting substitution. This may be due to its higher viscosity affecting efficient mixing or perhaps another property not considered here. In propylene carbonate, the fluoride leaving group, which is normally a very poor nucleophile in solution,[34] may be now strong enough to perform the reverse reaction or at least compete with the thiolate. In general, we observed that selective high conversion to disubstitution or tetrasubstitution could be controlled by the choice of solvent-base combinations that is, 1,4-disubstitution can be obtained cleanly using DMSO/DIPEA or THF/Cs_2CO_3 or propylene carbonate/DBU. Interestingly, DMF (used in perfluoroaryl peptide stapling, Table 1) does not appear to be optimal for this transformation, especially if in combination with DIPEA; and MeCN/DBU could be a valuable alternative. However, while 1,2,4,5-tetrasubstitution could be obtained using DMSO-DBU/Cs_2CO_3, this is only afforded as a moderate product (approx. 40%) and further optimisation was required.

3.1.1 | Two-level FD—screening experiment

The above analyses indicated that the reaction outcome may be sensitive to thiol concentration and the nature of the base and solvent. A more detailed FD screening experiment was used to further explore the combined roles of reaction time, temperature, concentration and reagent molar equivalence on product outcome and that may be used to control di- and tetra- fluoride-thiol substitution. This employed a two-level, five factorial full design space (Table 2) varying reaction time, temperature, solvent volume, molar equivalents of thiol and molar equivalents of base. A DMSO/Cs_2CO_3 system was selected as the combination of a polar aprotic solvent and a strong base (e.g., Cs_2CO_3, K_2CO_3 or DBU) increased the formation of substitution products; and the selection of a highly polar solvent would enable all reagents and products to remain in solution, which would expedite reaction profiling through ^{19}F NMR.

Specific combinations of each factor were combined, such that the relationship between each factor and the formation of specific substitution products (**3** or **4**) could be established with relatively few individual experiments and using simple statistical analyses. A total of 32 individual experiments were performed in disposable fritted syringes to minimize vessel headspace to avoid possible evaporation of HFB. After the completion of the reaction time (4 or 168 hours), samples were filtered, removing cesium salts and unreacted cesium carbonate and the reaction outcomes were measured using ^{19}F NMR (Figure 2A). In each case, only the major products or starting material were integrated, yet in some cases, some other minor ^{19}F resonances were observed (mostly mono-substitution but some others unidentified were generally observed with higher temperature and longer reaction time for example, reactions 9 and 31—see Supporting Information for raw data (Table S1) and ^{19}F NMR spectra (Figures S5-S35).

Analysis of variance (ANOVA, Figure 2B) indicated that the major factors that influence the product outcome are time, base and combinations of base-time and thiol-temperature; and with *P*-values <.05, these are unlikely to be random observations. Main effects analysis (Figure 2C,D) also indicated that both the amount of base and the time allotted for the reaction exert significant influence on the substitution outcome, as indicated by the steep gradient. It was clear that a greater amount of base generally afforded a higher proportion of tetrasubstitution over disubstitution, while longer reaction times also tended to afford higher conversion to higher substituted products. Unsurprisingly, the exact opposite was true for each factor in

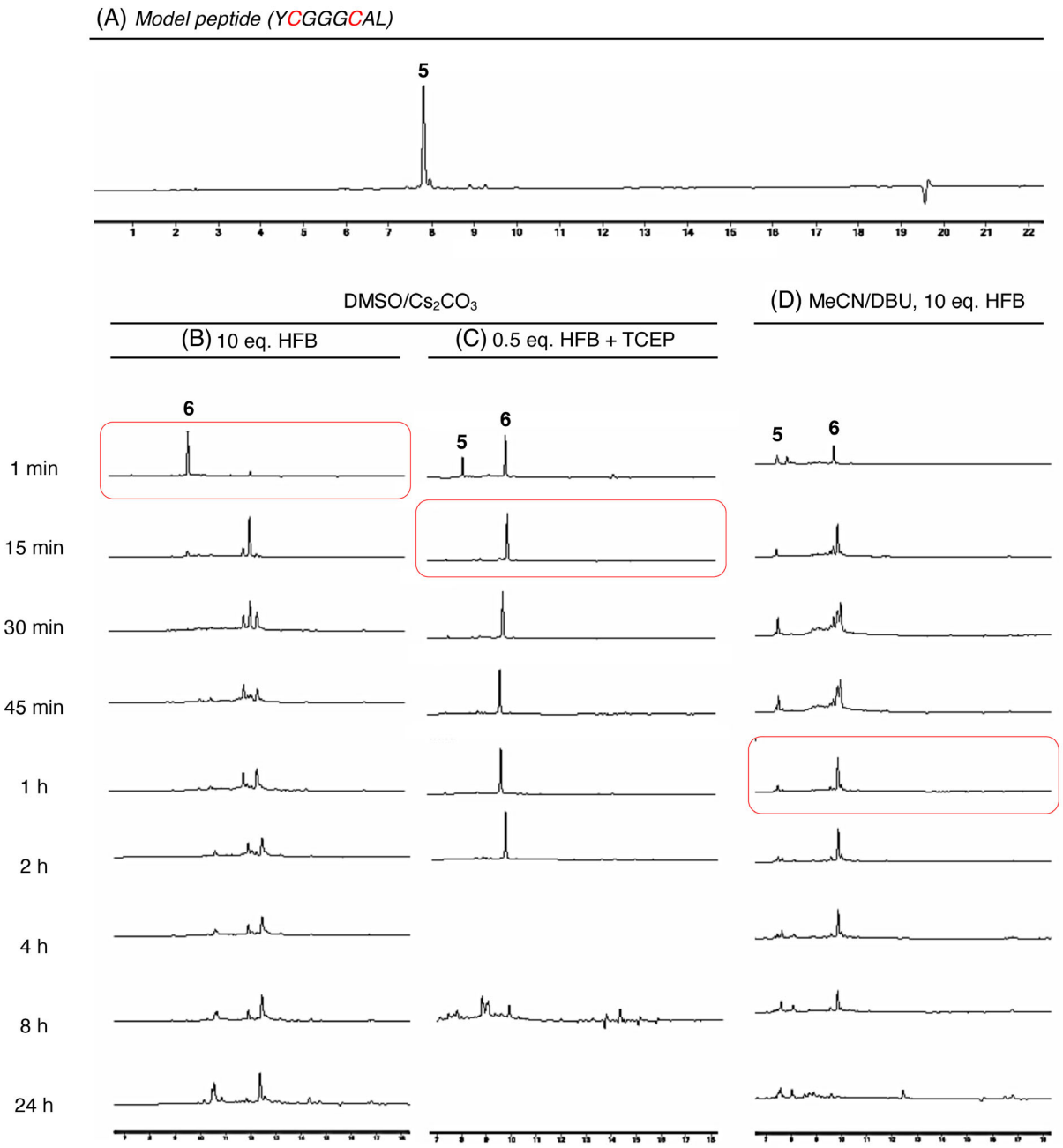


FIGURE 3 UV chromatograms (response at λ = 280 nm vs acquisition time in minutes) from LCMS analysis of: A, the model peptide (**5**) and reactions B-D sampled after 1, 15, 30, 45 minutes, 1, 2, 4, 8 and 24 hours under different conditions. The red boxes highlight the first timepoints for each condition where complete conversion of the linear peptide (**5**) into the stapled peptide (**6**) was observed

promoting 1,4-disubstitution and affords a means of discrimination between a disubstituted (stapled/cyclic) peptide and higher-substitution. Perhaps surprisingly, on average, the molar equivalents of thiol, reaction temperature and reaction concentration had comparatively lower influences over product selectivity and these variables *alone* were not a strong indicator of reaction outcome. In fact, there are examples of both high and low values for thiol, temperature and volume in both disubstitution favoring and tetrasubstitution favoring reactions. This is not to suggest that such factors do not influence the reaction, but likely indicates a more complex interaction with another factor (Figures S36-S38) that together have a more profound influence, which underlines the value of using a multifactorial analysis approach.

3.2 | Perfluoroaryl peptide-stapling

These outcomes of the base-solvent combination screenings and FD studies were applied to the exploration of cysteine-containing peptide stapling around HFB. Despite this transformation being typically

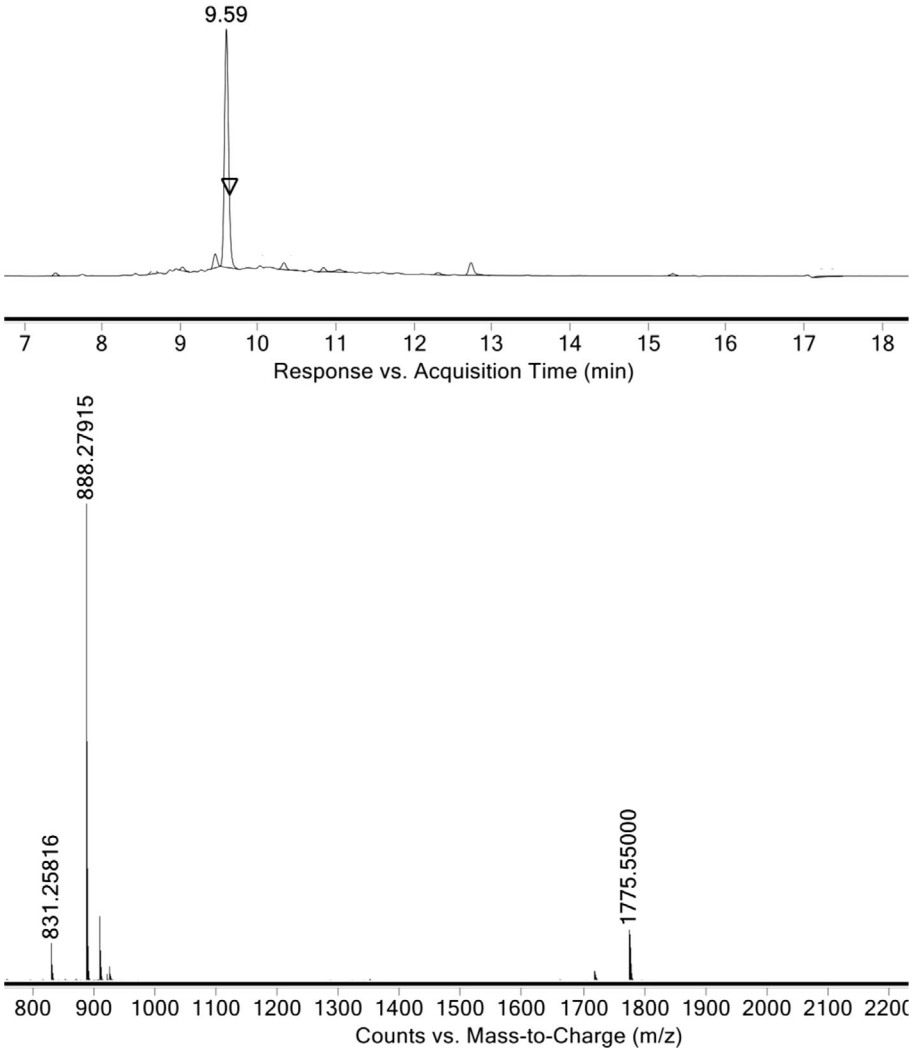


FIGURE 4 UV chromatogram (λ = 280 nm) and mass spectrum (ESI) of the crude cleavage product following on-resin perfluoroaryl-stapling reaction with DMF/DIPEA

performed using DMF (Table 1), our exploration indicated that MeCN and DMSO may provide the best combination of solubility (peptide and base) and potential for clean and rapid reaction (Table 4).

As such, the combinations MeCN/DBU and DMSO/Cs_2CO_3 were applied to a model di-cysteine peptide system to explore their application in peptide stapling and cyclisation (Scheme 3). Due to the increased complexity of the peptide reaction compared to the model reaction, and because ^{19}F NMR peak intensities are affected by the high amount of residual HFB (in the reaction mixture) and TFA (from peptide cleavage), LCMS analysis was used as the main tool for reaction monitoring and product characterization. By sampling the reaction mixtures at different time points, we produced a kinetic profile of the product formation (Figure 3B-D). In the reaction performed in DMSO and using Cs_2CO_3, the complete conversion of the linear peptide (**5**) into the stapled form (**6**) occurred almost instantaneous (<1 minute, Figure 3 B). Nevertheless, it appeared that little of **6** remained after 15 minutes under these conditions. Performing the same reaction in MeCN and using DBU as base gave a slower but complete conversion of the starting peptide into a cleaner stapled product (within approx. 1 hour, Figure 3D). These reactions appear to happen faster than those of the model system with N-acetyl cysteine and this is possibly a proximity effect (second substitution step is intramolecular).

In each case (Figure 3B,D), the presumed degradation products could not be ascribed to any identifiable products and no higher order substitution (i.e., tetrasubstitution - "double stapling") products were observed after prolonged reaction times. We hypothesized that a larger amount of peptide might favor the formation of such multicyclic systems. Therefore, the fast and clean reaction in DMSO/Cs_2CO_3 was repeated with a lower amount of HFB (around 0.5 eq., Figure 3C) and TCEP. We could verify that the combination of these factors afforded the stapled product rapidly and cleanly as before, but in this case, **6** remained detectable/stable for at least 2 hours and no tetra-substitution products were observed. The reaction was complete even with a stoichiometric or sub-stoichiometric amount of HFB and TCEP may play and important role in maintaining the stability of the product.

There are many examples of single-component stapling/cyclisation (e.g., alkene metathesis, lactamisation) carried out on solid-supported peptides, whereas similar two-component reactions are generally performed in solution because of potential site-isolation and by-products formation on resin.[39] We also wanted to determine whether the perfluoroaryl-stapling of a di-cysteine peptide was

possible on resin. Solid phase reactions offer undeniable advantages such as user-friendly handling, no laborious workup, easier purification and possible automation, resulting in a greener, faster, cheaper and possibly higher yielding process. Unfortunately, our optimal stapling conditions were deemed unsuitable for polystyrene resins, therefore, we investigated the efficiency of the standard solution-phase procedure (DMF, DIPEA) for stapling a 2-Cys peptide on-resin.[26,27] After selective on-resin trityl deprotection of cysteine residues, the stapled product (**6**) was successfully obtained after overnight shaking of the resin with a reaction mixture made of HFB and DIPEA in DMF (Figure 4).

Finally, to demonstrate that the HFB-mediated cyclisation on solid-phase could be applied to a longer and more complex peptide sequence, we successfully cyclised the Skin Penetrating and Cell Entering (SPACE) disulfide peptide[40,41] (peptide **7**) on-resin using DMF/DIPEA. The cyclised product (peptide **8**) was obtained with clean conversion and only a small amount of starting material remaining after 18 hours reaction time and following cleavage from the resin (Scheme S1 and Figures S43, S44). In addition to the procedures reported in Table 1, this is further evidence that this peptide stapling technique is applicable to changes in the peptide sequence components, length or inter-thiol spacing.

4 | CONCLUSIONS

In conclusion, the application of a two-level FD study in combination with ^{19}F NMR, has provided a more detailed understanding of the reactivity of HFB toward N-acetylcysteine and how this is particularly sensitive to base and solvent effects. This work also provided new conditions that afford the selective preparation of either 1,4-di *or* 1,2,4,5-tetra thiol-fluoride substitution products, principally controlled by the above factors. It is envisaged that these approaches can be exploited in future for the synthesis of branched or multicyclic peptide systems and poststapling modifications. Finally, new conditions (DMSO/Cs_2CO_3 and MeCN/DBU) that permit rapid (<1 minute and < 1 hour, respectively), clean and selective peptide stapling under peptide-compatible conditions were introduced. The products obtained using our procedure (without purification) are of equal if not better crude purity than previously reported peptide stapling approaches.[15,26,27] We also demonstrate that the 2-component on-resin perfluoroaryl stapling is achievable with high crude conversion using conditions that were previously only used in solution.

ACKNOWLEDGMENT

This work was supported by an EPSRC First Grant (EP/R020299/1) that employed PK and a University Alliance Doctoral Training Alliance COFUND/Marie Skłodowska-Curie (European Union's Horizon 2020 research and innovation programme, grant agreement No 801604) PhD Fellowship Programme in Applied Biosciences for Health (PD).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ORCID

Christopher R. Coxon https://orcid.org/0000-0002-3375-3901

REFERENCES

[1] N. N. Vorozhtsov, V. E. Platonov, G. G. Yakobson, *Bull. Acad. Sci.* **1963**, *12*, 1389. https://doi.org/10.1007/BF00847820.
[2] R. P. Mason, W. Rodbumrung, P. P. Antich, *NMR Biomed.* **1996**, *9*(3), 125. https://doi.org/10.1002/(SICI)1099-1492(199605)9:3<125::AID-NBM405>3.0.CO;2-F.
[3] A. Lattanzi, C. De Fusco, A. Russo, A. Poater, L. Cavallo, *Chem. Commun.* **2012**, *48*(11), 1650. https://doi.org/10.1039/c2cc17488j.
[4] L. A. Wall, W. J. Pummer, J. E. Fearn, J. M. Antonucci, *J. Res. Natl. Bur. Stand. Sect. A Phys. Chem.* **1963**, *67A*(5), 481. https://doi.org/10.6028/jres.067a.050.
[5] A. J. J. Lennox, *Angew. Chem. Int. Ed.* **2018**, *57*(45), 14686. https://doi.org/10.1002/anie.201809606.
[6] S. E. Denmark, A. J. Cresswell, *J. Org. Chem.* **2013**, *78*(24), 12593. https://doi.org/10.1021/jo402246h.
[7] I. P. Beletskaya, G. A. Artamkina, A. Y. Mil'Chenko, P. K. Sazonov, M. M. Shtern, *J. Phys. Org. Chem.* **1996**, *9*(6), 319. https://doi.org/10.1002/(SICI)1099-1395(199606)9:6<319::AID-POC786>3.0.CO;2-7.
[8] H. Wu, W. Chi, G. Baryshnikov, B. Wu, Y. Gong, D. Zheng, X. Li, Y. Zhao, X. Liu, H. Ågren, L. Zhu, *Angew. Chem. Int. Ed.* **2019**, *58*(13), 4328. https://doi.org/10.1002/anie.201900703.
[9] L. Testaferri, M. Tingoli, M. Tiecco, *J. Org. Chem.* **1980**, *45*(22), 4376. https://doi.org/10.1021/jo01310a022.
[10] B. Gierczyk, G. Schroeder, *Mendeleev Commun.* **2009**, *19*(2), 75. https://doi.org/10.1016/j.mencom.2009.03.007.
[11] M. Villa, M. Roy, G. Bergamini, M. Gingras, P. Ceroni, *Dalton Trans.* **2019**, *48*(12), 3815. https://doi.org/10.1039/c9dt00251k.
[12] J. N. Lowe, D. A. Fulton, S. H. Chiu, A. M. Elizarov, S. J. Cantrill, S. J. Rowan, J. F. Stoddart, *J. Org. Chem.* **2004**, *69*(13), 4390. https://doi.org/10.1021/jo030283o.
[13] J. Ponce González, M. Edgar, M. R. J. Elsegood, G. W. Weaver, *Org. Biomol. Chem.* **2011**, *9*(7), 2294. https://doi.org/10.1039/c0ob00790k.
[14] O. Jacobson, X. Yan, Y. Ma, G. Niu, D. O. Kiesewetter, X. Chen, *Bioconjug. Chem.* **2015**, *26*(10), 2016. https://doi.org/10.1021/acs.bioconjchem.5b00278.
[15] A. M. Spokoyny, Y. Zou, J. J. Ling, H. Yu, Y. S. Lin, B. L. Pentelute, *J. Am. Chem. Soc.* **2013**, *135*(16), 5946. https://doi.org/10.1021/ja400119t.
[16] S. Niembro, A. Vallribera, M. Moreno-Mañas, *New J. Chem.* **2008**, *32*(1), 94. https://doi.org/10.1039/b707776a.
[17] R. D. Kennedy, C. W. MacHan, C. M. McGuirk, M. S. Rosen, C. L. Stern, A. A. Sarjeant, C. A. Mirkin, *Inorg. Chem.* **2013**, *52*(10), 5876. https://doi.org/10.1021/ic302855f.
[18] T. Umemoto, L. M. Garrick, N. Saito, *Beilstein J. Org. Chem.* **2012**, *8*(1), 461. https://doi.org/10.3762/bjoc.8.53.
[19] M. Arisawa, T. Suzuki, T. Ishikawa, M. Yamaguchi, *J. Am. Chem. Soc.* **2008**, *130*(37), 12214. https://doi.org/10.1021/ja8049996.
[20] M. Mayor, J.-M. Lehn, *Helv. Chim. Acta* **1997**, *80*(8), 2277. https://doi.org/10.1002/hlca.19970800802.
[21] Y. Zou, A. M. Spokoyny, C. Zhang, M. D. Simon, H. Yu, Y. S. Lin, B. L. Pentelute, *Org. Biomol. Chem.* **2014**, *12*(4), 566. https://doi.org/10.1039/c3ob42168f.
[22] C. Zhang, P. Dai, A. M. Spokoyny, B. L. Pentelute, *Org. Lett.* **2014**, *16*(14), 3652. https://doi.org/10.1021/ol501609y.
[23] R. J. Steel, M. A. O'Connell, M. Searcey, *Bioorg. Med. Chem. Lett.* **2018**, *28*(16), 2728. https://doi.org/10.1016/j.bmcl.2018.03.003.

[24] J. E. Swedberg, C. I. Schroeder, J. M. Mitchell, T. Durek, D. P. Fairlie, D. J. Edmonds, D. A. Griffith, R. B. Ruggeri, D. R. Derksen, P. M. Loria, S. Liras, D. A. Price, D. J. Craik, *Eur. J. Med. Chem.* **2015**, *103*, 175. https://doi.org/10.1016/j.ejmech.2015.08.046.
[25] A. C. Conibear, S. Chaousis, T. Durek, K. Johan Rosengren, D. J. Craik, C. I. Schroeder, *Biopolymers* **2016**, *106*(1), 89. https://doi.org/10.1002/bip.22767.
[26] D. Gimenez, C. A. Mooney, A. Dose, G. Sandford, C. R. Coxon, S. L. Cobb, *Org. Biomol. Chem.* **2017**, *15*(19), 4086. https://doi.org/10.1039/c7ob00283a.
[27] S. J. M. Verhoork, C. E. Jennings, N. Rozatian, J. Reeks, J. Meng, E. K. Corlett, F. Bunglawala, M. E. M. Noble, A. G. Leach, C. R. Coxon, *Chem. - A. Eur. J.* **2019**, *25*(1), 177. https://doi.org/10.1002/chem.201804163.
[28] Y. Yin, Q. Fei, W. Liu, Z. Li, H. Suga, C. Wu, *Angew. Chem. Int. Ed.* **2019**, *58*(15), 4880. https://doi.org/10.1002/anie.201813827.
[29] R. Leardi, *Anal. Chim. Acta* **2009**, *652*(1–2), 161. https://doi.org/10.1016/j.aca.2009.06.015.
[30] E. Marcucci, N. Bayó-Puxan, J. Tulla-Puche, J. Spengler, F. Albericio, *J. Comb. Chem.* **2008**, *10*(1), 69. https://doi.org/10.1021/cc7001588.
[31] D. Gimenez, A. Dose, N. L. Robson, G. Sandford, S. L. Cobb, C. R. Coxon, *Org. Biomol. Chem.* **2017**, *15*(19), 4081. https://doi.org/10.1039/c7ob00295e.
[32] S. B. Lawrenson, R. Arav, M. North, *Green Chem.* **2017**, *19*, 1685. https://doi.org/10.1039/c7gc00247e.
[33] M. Gazitúa, R. A. Tapia, R. Contreras, P. R. Campodónico, *New J. Chem.* **2018**, *42*(1), 260. https://doi.org/10.1039/c7nj03212a.
[34] V. M. Vlasov, *J. Fluor. Chem.* **1993**, *61*, 193. https://doi.org/10.1016/S0022-1139(00)80104-9.
[35] D. Perrin, *Dissociation Constants of Organic Bases in Aqueous Solution*, Butterworths, London **1965**.
[36] F. Ravalico, S. L. James, J. S. Vyle, *Green Chem.* **2011**, *13*(7), 1778. https://doi.org/10.1039/c1gc15131b.
[37] R. Srivastava, *J. Mol. Catal. A Chem.* **2007**, *264*(1–2), 146. https://doi.org/10.1016/j.molcata.2006.09.012.
[38] D. Perrin, *Dissociation Constants of Inorganic Acids and Bases in Aqueous Solution*, Butterworths, London **1969**.
[39] Y. H. Lau, P. De Andrade, Y. Wu, D. R. Spring, *Chem. Soc. Rev.* **2015**, *44*(1), 91. https://doi.org/10.1039/c4cs00246f.
[40] T. Hsu, S. Mitragotri, *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*(38), 15816. https://doi.org/10.1073/pnas.1016152108.
[41] M. Chen, M. Zakrewsky, V. Gupta, A. C. Anselmo, D. H. Slee, J. A. Muraski, S. Mitragotri, *J. Control. Release* **2014**, *179*(1), 33. https://doi.org/10.1016/j.jconrel.2014.01.006.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Dognini P, Killoran PM, Hanson GS, et al. Using ^{19}F NMR and two-level factorial design to explore thiol-fluoride substitution in hexafluorobenzene and its application in peptide stapling and cyclisation. *Peptide Science*. 2020;e24182. https://doi.org/10.1002/pep2.24182